DOCUMENT RESUME

ED 204 414 TM 810 494

AUTHOR Covert, Robert W.; Fang, Wei Li
TITLE The Joint Dissemination Review Panel: Can Approved Submittals Be Distinguished from the Rejections?
PUB DATE Apr 81
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Demonstration Programs; Educational Practices; *Evaluation Criteria; *Evaluation Methods; Information Dissemination; Instructional Materials; Program Effectiveness; *Program Validation; Reliability; *Validated Programs
IDENTIFIERS *Joint Dissemination Review Panel

ABSTRACT

The Joint Dissemination Review Panel (JDRP) examines educational interventions and determines their effectiveness. This study's purpose was to assess this review process and identify common characteristics of submittals and ways in which approvals and rejections differed. A locally-developed instrument which collected data on the intervention's effectiveness was administered to 96 submittals reviewed during 1978 and 1979. Results indicated that submittals contained measurable objectives, evaluation designs, and specified sample sizes. Using tests of significance, approved submittals significantly differed from rejections in that they used multiple testing measures; results were statistically significant; information was provided on salient features; and the narrative and tables were clearly presented. Discriminant analysis demonstrated that approvals reported using quasi-experimental designs, multiple test procedures, more information on key features, clearer tables, and control of selection and regression effects. One variable, statistically significant test results, was inversely related. Based on these findings, JDRP submittals should include measurable objectives, quasi- or experimental designs, multiple testing measures, reliable and valid tests, statistically significant test results, a clearly described intervention, and replication information. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED204414

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✗ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

The Joint Dissemination Review Panel: Can Approved Submittals Be Distinguished from the Rejections?

Paper Presentation
for
Evaluation--Implications for Policy,
Annual Meeting of the American Educational Research Association
Los Angeles, California
April 13-17, 1981

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
W. L. Fang
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Prepared by:

Robert W. Covert, Ph.D.
Wei Li Fang, Ph.D.
Evaluation Research Ctr.
University of Va.
164 Rugby Rd.
Charlottesville, Va. 22903
(804) 924-7163

TM 810 464

The Joint Dissemination Review Panel (JDRP) is a federal review board which examines educational products and practices and determines whether the submitting project provided persuasive evidence of effectiveness. Effectiveness criteria state that there is evidence of positive impact, that effects are statistically significant and educationally important, that measures are reliable and valid, that the observed effects can be attributed to the intervention, that the evidence is credible, and that the intervention and its effects can be replicated in other sites. If JDRP review results in approval, the intervention is recognized as having the capability of producing its claimed effects. The intervention may then be nationally disseminated. Thus the Panel serves as a screening mechanism for national dissemination through a review process that substantiates the validity of the project's claims.

The purpose of the present study was to examine the review process based on the identified dimensions of evidence of effectiveness. Three research questions were asked:

1. What were the characteristics of submittals reviewed by JDRP?

2. Which independent factors differentiated between approved and rejected submittals?

3. What set of variables discriminated between approved and rejected submittals?

To answer these questions, an instrument was devised that collected data related to the submitting project's demographic characteristics, the

nature of the educational intervention, and evidence of effectiveness. Development of the instrument was based on documents related to the effectiveness criteria and the review process (Education Division, Note 1; National Testing Service Research Corporation, Note 2; Network, Note 3; Pyecha & Fisher, Note 4; Tallmadge, Note 5) and evaluation literature (Bernstein, Bohrnstedt, & Borgatta, 1975; Bracht & Glass, 1968; Campbell, 1969; Campbell & Stanley, 1963; Cook & Campbell, 1979).

The sample consisted of 83 projects that submitted their interventions for review by the Panel during 1978 and 1979. Of these 83 projects, 70 submitted on only one occasion while the remaining 13 submitted twice. In the case of the projects that submitted only once, 52 submittals were approved, and 18 were rejected. In the case of the 13 projects that submitted twice, the first submittal was initially rejected while the resubmission was approved. Since these projects appeared to be two distinct groups, they were examined as separate sub-samples.

The locally-developed instrument was field tested prior to administration to each of the 96 submittals. Using ten randomly selected submittals, three reviewers independently read each submittal and recorded the appropriate data on the instrument. Rater agreement was 81% on items related to the reporting of evidence of effectiveness.

Upon review of each of the 96 submittals, information on each item on the instrument was then collected by the investigator. Since a number of the submittals reported data on more than one objective, a decision was made to assess only those objectives that were cognitive in nature. Succeeding items then collected information which corresponded to the criteria employed by JDRP as they related to the selected cognitive objective.

These items elicited the following types of data:

	Variable (Operational Definition)	JDRP Criterion
1.	evaluation design	evidence of impact
2.	sampling procedures	credibility of evidence
3.	type of comparison group	evidence of impact
4.	statistical significance	evidence of statistical reliability of effects
5.	educational importance	evidence that the effects are educationally meaningful
6.	internal validity	evidence that the effects are attributable to the intervention
7.	external validity	evidence of generalizability
8.	replication	evidence of generalizability
9.	type of instrument either standardized or locally-developed	interpretability of measures
10.	reliability of the instrument	intepretability of measures
11.	validity of the instrument	interpretability of measures
12.	test administration procedures	credibility of evidence
13.	types of reported scores	credibility of evidence
14.	types of data analysis procedures	credibility of evidence

To answer research question 1, frequency information was tabulated on each of the instrument's items. These data were collected on both sub-samples: those that submitted only once and those that submitted twice. Results indicated that each of the 96 submittals contained measurable objectives, an evaluation design, and a specified sample size and administered at least one instrument. Inclusion of data related to the identified evidence of effectiveness criteria varied depending on the particular item. Tables 1,2, and 3 present a summary of the information that submittals contained on selected variables related to evidence of effectiveness.

Table 1

Number of Submittals that Included Data on Selected Variables

Variable	S1 Submittals (n = 70) n	%	S2 Rejections (n = 13) n	%	S2 Approvals (n = 13) n	%
Measurable objectives	70	100	13	100	13	100
Quasi- or experimental design	34	49	6	46	7	53
Comparison standard[a]	64	91	11	85	10	76
Representative sample[b]	18	26	2	16	5	38
History threat controlled	45	64	8	62	9	69
Maturation threat controlled	41	59	8	62	8	62
Testing threat controlled	45	64	9	69	9	69
Instrumentation threat controlled	43	61	9	69	9	69
Regression threat controlled	19	27	2	15	4	31
Selection threat controlled	53	76	10	77	11	85
Mortality threat controlled	40	57	3	23	5	38
Interaction with selection threat controlled	8	11	1	8	2	15
Ambiguity threat controlled	53	76	4	31	13	100
Rivalry threat controlled	5	7			1	8
Equalization threat controlled	2	3				
Demoralization threat controlled	3	4				
Diffusion threat controlled	2	3				
Experimenter effect threat controlled	17	24	3	23	4	31
Hawthorne threat controlled	4	6				
Setting interaction threat controlled	1	1				
Selection and treatment interaction controlled	54	77	11	85	11	85
One or more instruments administered	70	100	13	100	13	100
Educational importance (s.d. formula)[c]	35	50	3	23	7	54
Replication	43	61	7	54	9	69
Cost information included	37	53	7	54	7	54

[a]Figures included those projects that used a national normative group.

[b]Figures included only those projects that used either probability sampling or a combination of probability-nonprobability procedures in obtaining the treatment group.

[c]Figures included those projects that provided information about the intervention's educational importance for some or all test results.

[d]Figures included only those projects that provided information on estimated costs for installation and operation.

Table 2

Number of Submittals that Included Data
on Variables Related to Standardized Tests

Variable	S1 Submittals (n = 40) n	%	S2 Rejections (n = 9) n	%	S2 Approvals (n = 8) n	%
Followed publisher's guidelines	35	88	0		0	
Reported all test results	10	25	6	67	8	100
Statistical significance of reported test results[a]	21	53	4	44	6	75
Reported test statistic	28	70	7	78	7	88
Inconsistencies in reporting	9	23	7	78	5	63

[a]Figures included those projects that provided data related to the statistical significance of some or all test results.

Table 3

Number of Submittals that Included Data
on Variables Related to Locally-developed Tests

Variable	S1 Submittals (n = 33) n	%	S2 Rejections (n = 8) n	%	S2 Approvals (n = 7) n	%
Used standardized test administration procedures	13	39	0		0	
Reliability of test reported	23	70	5	62	5	71
Validity of test reported	19	58	2	25	3	41
Reported all test results	28	85	8	100	7	100
Statistical significance of reported test results[a]	18	55	2	25	5	71
Reported test statistic	26	65	8	100	7	100
Inconsistencies in reporting	12	36	2	25	2	29

[a]Figures included those projects that provided data related to the statistical significance of some or all test results.

In the two sub-samples approved and rejected submittals were compared to obtain the answer to research question 2. Chi-square tests of significance and t-tests were employed to determine whether there were statistically significant differences. In the sub-sample where projects submitted on only one occasion, results showed that approved submittals significantly differed from rejected submittals in the following ways:

1. they reported using a higher number of testing measures;
2. information was provided on the salient features of the intervention;
3. the educational importance of the intervention was discussed in the narrative; and
4. the information in the tables was clearly presented.

In the sub-sample where projects submitted on two occasions, findings indicated that approved submittals significantly differed from rejected submittals in the following ways:

1. they reported controlling a higher number of internal validity threats;
2. they reported using a higher number of testing measures;
3. reported results were statistically significant;
4. information was provided on the salient features of the intervention; and
5. the information in the narrative and tables was clearly presented.

In both sub-samples, approved submittals tended to provide information related to the educational importance of the reported test results.

Educational importance was defined using the JDRP-sanctioned standard deviation formula, $(\overline{Y}_t - \overline{Y}_c) \geq (\sigma/2$ of the normative group). Tables 4 and 5 present the statistics on those variables in which statistical significance was found.

Table 4

Statistically Significant Differences
That Were Found Between Approved and Rejected Submittals
That Were Reviewed on Only One Occasion

Variable	Scale Values	Rejection $\overline{X}$	Approval $\overline{X}$	Statistic	df	P
Number of test measures	1-28	1.2	2.5	$t = 2.4$	54	.02
Salient features	1-3	2.2	2.5	$t = 3.0$	42	.01
Educational importance discussed in narrative	1-2	1.5	1.8	$\chi^2 = 4.9$	1	.03
Educational importance (s.d. formula)	1-3	1.4	1.8	$t = 1.9$	68	.06
Clear tables	1-3	1.6	2.3	$t = 3.3$	68	.002

Table 5

Statistically Significant Differences
That Were Found Between Approved and Rejected Submittals
That Were Reviewed on Two Occasions

Variable	Scale Values	Rejected $\bar{X}$	Approval $\bar{X}$	Statistic	df	p
Number of controlled internal validity threats	1-13	4.2	5.5	t= 6.6	25	.001
Number of test measures	1-28	1.5	1.9	t= 2.0	25	.05
Statistical significance	1-3	2.1	2.8	t= 4.5	25	.001
Salient features	1-3	1.9	2.5	t= 4.9	25	.001
Educational importance (s.d. formula)	1-3	1.2	1.8	t= 2.1	12	.06
Clear narrative	1-5	2.5	3.5	t= 5.9	25	.001
Clear tables	1-3	1.3	2.4	t= 2.1	25	.05

To answer research question 3, a discriminant analysis procedure was performed on variables related to JDRP review criteria. Only projects that submitted once to JDRP were considered in the analysis. Analysis demonstrated that six variables were directly related to the approved submittals. These approvals reported the use of quasi-experimental or experimental designs more frequently than pre-experimental designs, a higher number of tests, a greater amount of information on the intervention's salient features, clearer tables, and the elimination of selection and statistical regression effects. Only one variable, the statistical significance of test results, was inversely related to the function. These results provided additional evidence that approvals could be distinguished from rejections on selected variables related to the intervention's effectiveness. Table 6 presents these data.

Table 6

Variables that Account for the Function
Using Step-wise Discriminant Analysis Procedures

Predictor Variable	Scale Values	Discriminant Function Coefficient	Wilk's Lambda	Level of Significance
Clarity of tables	1-3	-.59	.86	.002
Selection threat controlled	1-2	-.60	.78	.001
Regression threat controlled	1-2	-.49	.74	.001
Information of salient features	1-3	-.53	.70	.001
Evaluation design	1-2	-.33	.67	.001
Statistical significance of reported tests	1-3	.35	.65	.001
Number of instruments	1-28	-.31	.63	.001

Note. Percent of explained variance was 37.5%. Group $\overline{X}$ for approvals was -.44; group $\overline{X}$ for rejections was 1.30.

The results of this study have implications for three types of audiences--program developers/ demonstrators, potential adopters/adapters at the local level, and JDRP members. In terms of the program developers who are considering submission to JDRP, the study has demonstrated that all submittals, regardless of the final Panel decision, should contain a statement of the objectives written in measurable terms, an evaluation design that utilizes a comparison standard and a specified sample size for the treatment group, and should provide evidence that at least one instrument was employed to measure the intervention's effects. The use of multiple measures is recommended since findings indicate that the higher the number of testing measures that were implemented, the greater the chances of approval. Results also emphasize the need for projects to provide information on the key features of the educational practice or product. A clear description of the essential intervention components enables local educators to make a decision about adopting a particular intervention. The submittal should also contain a discussion of the educational importance of the practice or product. Program developers are also cautioned to select evaluation designs in which the effects are clearly attributable to the intervention. Approved submittals tended to show evidence that the third-variable threats of selection, history, and statistical regression were controlled. Although there was no statistical significance between approved and rejected submittals on the type of evaluation design (i.e., pre-experimental, quasi-experimental, experimental) that was employed, in interactions with panel members several noted a

personal bias toward the use of experimental designs. On the other hand, the Education Division (Note 1) and other Panel members have stated that an experimental design is not necessary as long as data are presented that show a comparison standard was utilized.

Findings demonstrate that the type of testing measure (i.e., standardized vs. locally-developed) that is employed does not affect the JDRP decision. However, if the submitter decides to use a locally-developed test, it is advisable to report reliability and validity information. Although no statistically significant differences were found on this variable, both JDRP submittal guidelines and several Panel members emphasize this point.

No statistically significant differences were obtained when the investigator looked at the type of data that were reported for each of the testing measures in the submittals. However, when one reviews the data related to implementation, analysis, and reporting variables (see Tables 1,2,and 3), it is evident that projects in the sub-sample that submitted on two occasions included more data in their resubmittals. This is particularly true in terms of the test results; the test statistic; and information on replication.

In terms of the local practitioners who are considering the adoption of a JDRP-approved educational practice or product, they will generally find that information related to measurable objectives, the use of a comparison standard, the reporting of reliable and valid measures, and the statistical significance of the reported tests is usually available. Approximately 50% of the time, the intervention will have data related to the educational importance of the practice or product (i.e., in terms of the standard deviation formula); replication figures for more than one site

or time period; and cost effectiveness information. This cost effectiveness information may not give an accurate portrayal of the installation and recurring costs of the intervention. Accuracy will depend on the developer's reporting of the individual components that comprise each estimate.

Information concerning the control of internal and external validity threats will vary from submittal to submittal. Although selection effects and the ambiguity of the causal direction were frequently addressed (83% and 81% respectively), the elimination of other third-variable threats was not as high (see Table 1). The control of external validity threats was almost completely lacking. This is distressing when viewed from the perspective that these approvals are eligible for national dissemination and adoption.

Panel members may profit from this study in terms of seeing how an external evaluator has systematically examined a sample of the reviewed submittals. In interviews Panel members emphasized the importance of meeting each of the established review criteria, yet the data clearly demonstrate that all approvals do not meet each of these standards. This is particularly the case in the criteria of attribution (i.ei, the elimination of alternative explanations) and generalizability (i.e., the control of external validity threats). Of those variables that differentiated between the approvals and rejections, three were subjective measures -- the clarity of the tables, the amount of information presented on all salient features of the intervention, and the discussion of educational importance in the narrative. Perhaps additional variables are operating that would provide further clarification of the review process.

Findings demonstrated that there were a few single variables which distinguished between approvals and rejections. On the other hand, if one looked at a cluster of variables, it was found that a combination of seven variables discriminated between the approved and rejected submittals. These findings highlight the need for further studies in this area. A replication of the present study should be conducted, whereby the review criteria are verified across time and the use of the minutes from the review sessions is a part of the research procedures. Other areas of future study are the exploration of the feasibility of developing more detailed guidelines for potential submitters (e.g., presentation of cost effectiveness information, reporting of replication data) and the examination of the role of additional factors (e.g., Panel variables, Education Department reorganization) that may have a bearing on the review process.

Reference Notes

1. Education Division. Joint Dissemination Review Panel: Purpose, procedures, and criteria. March 1, 1977. (Available from Division of Educational Replication, U.S. Office of Education, 400 Maryland Avenue, S.W., Room B-732, TransPoint Building, Washington, D.C. 20202.)

2. National Testing Service Research Corporation. Joint Dissemination Review Panel (JDRP) technical assistance resource guide. Unpublished manuscript. (Available from NTS Research Corporation, 2634 Chapel Hill Boulevard, Durham, North Carolina 27707.)

3. Network. Program validation: Four case studies. 1978. (Available from the Network, 290 Aouth Main Street, Andover, Massachusetts 01810.)

4. Pyecha, J.N. & Fisher, M.D. Helpful hints for preparing and submitting ESEA, Title I regular and migrant education projects to the Education Division's Joint (NIE and USOE) Dissemination Review Panel. October 1978. (Available from Division of Educational Replication, U.S. Office of Education, 400 Maryland Avenue, S.W., Room B-732, TransPoint Building, Washington, D.C. 20202.)

5. Tallmadge, G.K. Ideabook: The Joint Dissemination Review Panel. Washington, D.C.: U.S.Office of Education, October 1977. (Available from Division of Educational Replication, U.S.Office of Education, 400 Maryland Avenue, S.W., Room B-732, TransPoint Building, Washington, D.C. 20202.)

References

Bernstein, I.H., Bohrnstedt, G.W., & Borgatta, E.F. Problems of external validity in evaluative research. Sociological Methods in Research, 1975, 4 (1), 101-128.

Bracht, R.F., & Wortman, P.M. The external validity of experiments. American Educational Research Journal, 1968, 5 (4), 437-474.

Campbell, D.T. Reforms as experiments. American Psychologist, 1969, 24, 409-429.

Campbell, D.T., & Stanley, J.C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Cook, T.D., & Campbell, D.T. Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally, 1979.